

Characterization of 14 anonymous nuclear loci in *Pinus thunbergii* and their cross-species transferability

Suharyanto • Mine Nose • Susumu Shiraishi

Received: 2010-12-07; Accepted: 2011-02-22

© Northeast Forestry University and Springer-Verlag Berlin Heidelberg 2012

Abstract: We characterized 14 anonymous nuclear loci from *Pinus thunbergii* Parl., an important pine species native to Japan. One hundred and twenty-six single nucleotide polymorphisms (SNPs) were identified from these loci, giving a frequency of 1 SNP per 51 bp. Nucleotide diversity (θ) ranged from 1.06×10^{-3} to 11.87×10^{-3} , with an average of 4.99×10^{-3} . Only one locus (mK45) deviated significantly from the Hardy-Weinberg equilibrium. Thirteen of 14 loci were applicable in other pine species. These loci will be useful for nucleotide variation studies and will provide material for SNP-based marker development in *P. thunbergii* and related species.

Keywords: anonymous nuclear loci; *Pinus thunbergii*; SNP; SCAR; cross-species transferability

Introduction

Pinus thunbergii Parl. is an important pine species, native to Japan, which is endangered because of pine wilt disease (Mamiya 1988). To cope with this problem, since 1978, a series of

genetic conservation and resistance breeding projects has focused on selection and creation of resistant individuals (Nose and Shiraishi 2008). To assist these projects, genetic studies using molecular markers are important.

Materials and methods

In this study, we developed 14 anonymous nuclear loci using a locus-specific sequence characterized amplified region (SCAR)-based approach (Paran and Michelmore 1993). The genomic library was obtained by screening the monomorphic band produced by 60 random amplified polymorphic DNA (RAPD) primers (13-mer) in 16 individuals of *P. thunbergii*. The RAPD-PCR (Williams et al. 1990) reaction was conducted in a total volume of 10 μ l containing 1 \times PCR buffer, 2.0 mM $MgCl_2$, 0.2 mM of each dNTP, 0.25 U Platinum *Taq* DNA polymerase (Invitrogen), 1.5 μ M primer, and 20 ng template DNA. The PCR cycles were as follows: 94°C for 1 min; followed by 40 cycles of 94°C for 30 s, 50°C for 30 s, and 72°C for 1 min; followed by a final extension of 72°C for 2 min. The amplification product was analyzed by 1.5% agarose gel electrophoresis. A monomorphic band with an approximate fragment length of 300–700 bp was excised from the gel and used as template DNA for re-amplification. The re-amplification product was ligated into the pGEM-T vector (Promega) and transformed into JM109 competent cells (Promega). Universal primers -21M13 and M13Rev were used to directly sequence both strands of the colony PCR product using BigDye Terminator v3.1 (Applied Biosystems) and electrophoresis on an ABI PRISM 3130 (Applied Biosystems). An internal locus-specific primer was designed from the sequencing results and used for PCR amplification and direct sequencing.

In order to screen for sequence polymorphisms, PCR products were amplified from between eight to sixteen individuals of *P. thunbergii* (Table 1). The PCR reaction was set up in a volume of 10 μ l containing 1 \times PCR buffer (Invitrogen), 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 0.25 U Platinum *Taq* DNA polymerase (Invitrogen), 0.2 μ M primer pair, and 20 ng template DNA. A

The online version is available at <http://www.springerlink.com>

Suharyanto  • Mine Nose

Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan.

E-mail: suharyantoster@gmail.com; Phone: (+81)80-5252-3435; Fax: (+81)92-642-2872

Mine Nose

Present address: Forest Tree Breeding Center, Forestry and Forest Products Research Institute, 3809-1 Ishi, Juo-cho, Hitachi, Ibaraki 319-1301, Japan. E-mail: mineppi@hotmail.com

Susumu Shiraishi

Faculty of Agriculture, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan E-mail: sushi@agr.kyushu-u.ac.jp; Phone and Fax: (+81)92-642-2872

Responsible editor: Chai Ruihai

touchdown PCR program was used in all primers as follows: 94°C for 1 min; followed by 10 touchdown cycles of 94°C for 30 s, 62–55°C (decreasing 0.7°C per cycle) for 30 s, and 72°C for 1 min; followed by 25 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 60 s; followed by a final extension at 72°C for 2 min. For the sequencing reaction, the same chemicals were used, as described above. Sequencing results were base-called and aligned using Sequencher v.4.2 (GeneCodes).

Homologous sites from each locus alignment that appeared to exhibit sequence variation, either in heterozygous or homozygous forms, were evaluated by manual inspection. Nucleotide diversity of each locus was quantified as $\theta = K/[L * [1^{-1} + 2^{-1} + 3^{-1} + \dots + (n-1)^{-1}]]$ (Watterson 1975; Primmer et al. 2002), where K is the number of observed polymorphic sites, L is the total length of the sequence (in bp), and n is the total number of chromosomes screened. Number of chromosomes screened is defined that in a diploid organism, it would carries two copies of

the locus, and thus two chromosomes were screened. The formula corrects for different sequence lengths and number of chromosomes screened among loci.

To test if the SNPs discovered in this study are useful for genetic studies, one of the most polymorphic sites in each locus was selected and used for estimation of the observed and expected heterozygosities (H_o , H_e) and tested for Hardy-Weinberg equilibrium. These analyses were done using Cervus v3.0.3 (Kalinowski et al. 2007) and Genepop online version (<http://genepop.curtin.edu.au/>), respectively.

Results and discussion

A list of anonymous nuclear loci generated in this study is presented in Table 1.

Table 1. Characterization of 14 anonymous nuclear loci, developed from *Pinus thunbergii*.

Locus	Primer sequence (5'-3')	No. of chromosomes screened	No. of bp screened	No. of SNPs	$\theta \times 10^3$	Heterozygosity			Accession no.
						SNP identity	H_o	H_e	
mK01	F: CTGAGTGAAGTTACGGTAGATTGAAG R: CGTGTCTTGTGTAAGCGATGTC	30	407	2	1.24	174: A/C	0.200	0.186	AB607011
mK03	F: CCCTGATAGCGATAAGCGATAAGAG R: CCTGATCCAACGGTCAATATGC	32	330	5	3.76	139: A/G	0.375	0.315	AB607012
mK08	F: TGCTTGAAAGAGCTCCATGTTC R: TTTTGTGTTGGTACAGTTCTTTCTCTG	32	618	3	1.21	310: C/T	0.313	0.272	AB607013
mK37b	F: CATGCTCTTGATTGGACTGCACAC R: GAAATGGAGGAGGATAAGGAGG	32	467	2	1.06	280: A/C	0.188	0.272	AB607016
mK45	F: GGTTGTGGACATTCAGTGAATTG R: TTGGTTGGCTCACGAAAGTTCTG	30	448	22	11.83	304: C/T*	0.000	0.515	AB607018
mK46a	F: GGTTGTGTTATTCATATCGCATG R: GGATGATACAAATGACAAGACCTCC	32	601	8	3.31	442: A/G	0.563	0.466	AB607019
mK48a	F: GTCTGGTTGTGTCAGGTGTGG R: ATCAAATTTTCTTACCTTGGCTACCTC	32	457	21	11.41	362: G/T	0.250	0.315	AB607020
mK49b	F: CAAGAAGAACTATGTGCCTCGTGATC R: GTGTCAAAGTATTACGATATTTCAAGTGG	32	481	24	11.87	159: A/T	0.063	0.175	AB607021
mK51a	F: CTCAAGCATAAGAGGTTGAAATCACC R: AATGCTCAGAGATATGCAATGAACC	16	479	3	1.89	308: A/C	0.000	0.233	AB607022
mK63a	F: TCCTGGTGGTATTCTACTCTTTTCG R: AAAATTGCAGCAAAAGTTGTGAAG	32	417	4	2.38	364: C/T	0.313	0.272	AB607023
mK65a	F: GACAAGCTCGAACCGACATCAA R: CGGTCATGTAATTTAATTGTGCTAAGC	32	405	3	1.84	255: C/T	0.063	0.063	AB607024
mK67b	F: TGTGCAATTTTCACTAGGAGTGTTG R: ACAAGCTCGATCCAGAGAGTTGG	32	340	16	11.69	131: A/C	0.063	0.175	AB607025
mK68b	F: GCTACTGTGGATGTAGCAGTTATGG R: AAGCTCGAACACAAAGATCG	32	443	3	1.68	148: A/G	0.438	0.353	AB607026
mK71b	F: AACCAATTGCTACCTTTATACCACTC R: CGACAAGCTCGTTCCATAACTCACTC	32	516	10	4.81	199: G/T	0.500	0.516	AB607027
Total/Average		428	6409	126	4.99				

θ , nucleotide diversity; SNP identity, position and alleles of SNP site used for calculation of observed heterozygosity (H_o), expected heterozygosity (H_e), and the goodness of fit to the Hardy-Weinberg equilibrium; *, significant deviation ($p < 0.05$).

Fourteen sets of locus-specific primers successfully produced a PCR product suitable for direct sequencing, i.e., the PCR amplification resulted in a single, strong band as visualized by agarose gel electrophoresis. High-quality sequences were obtained

from these loci with a total length of 6,409 bp. One hundred and twenty-six SNPs were identified, which translates to an average of 1 SNP per 51 bp. Nucleotide diversity (θ) of the individual loci ranged from 1.06×10^{-3} to 11.87×10^{-3} , with an average of

4.99×10^{-3} . The H_o ranged from 0.000 to 0.563, whereas H_e ranged from 0.063 to 0.516. Using only the most polymorphic SNP site within each locus, we found that one locus (mK45) violated the assumptions of the Hardy-Weinberg equilibrium ($p < 0.05$). The remaining 13 loci should be regarded as unlinked loci.

The developed loci were further tested in five pine species to confirm their cross-species applicability (Table 2).

Table 2. Cross-species transferability of 14 anonymous loci in 5 pine species. ++, single band in the expected size; +, band in expected size with other nonspecific bands; -, no PCR product.

Locus	Subsection <i>Pinus</i>			Subsection <i>Australes</i>	
	<i>Pinus luchuensis</i>	<i>Pinus densiflora</i>	<i>Pinus massoniana</i>	<i>Pinus palustris</i>	<i>Pinus taeda</i>
mK01	++	-	-	-	-
mK03	++	-	-	-	-
mK08	++	-	++	-	-
mK37b	++	++	+	-	-
mK45	-	-	-	-	-
mK46a	++	++	+	-	-
mK48a	++	++	++	++	++
mK49b	-	++	-	-	-
mK51a	++	++	++	++	++
mK63a	++	-	-	-	-
mK65a	++	++	++	-	-
mK67b	++	-	++	-	-
mK68b	++	++	++	-	-
mK71b	++	++	++	-	-

Four individuals from each pine species were used for the test. In 13 of 14 loci, a single PCR product of the expected size was successfully amplified in at least 1 species. In both *P. luchuensis*

and *P. densiflora*, seven loci were applicable, a result that was encouraging for further use of these loci in experiments to study the comparative nucleotide variation among these pines. The SNPs discovered in this study also provide important material for SNP-based marker development in *P. thunbergii* and related pine species.

Acknowledgments

We thank to the Fukuoka Forest Research and Extension Center for providing the needle materials.

References

- Kalinowski ST, Taper ML, Marshall TC. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, **16**: 1099–1006.
- Mamiya Y. 1988. Histories of pine wilt disease in Japan. *Journal of Nematology*, **20**: 219–226.
- Nose M, Shiraishi S. 2008. Breeding for resistance to pine wilt disease. In: Zhao BG, Futai K, Sutherland JR, Takeuchi Y. (eds), *Pine wilt disease*. Japan: Springer, pp. 334–350.
- Paran I, Michelmore RW. 1993. Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Theoretical and Applied Genetics*, **85**: 985–993.
- Primmer CR, Borge T, Lindell J, Saetre GP. 2002. Single-nucleotide polymorphism characterization in species with limited available sequence information: high nucleotide diversity revealed in the avian genome. *Molecular Ecology* **11**:603–612.
- Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, **6**: 217–250.
- Williams JGK, Kubelic AR, Livak KJ, Rafalski JA, Tingey SV. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*, **18**: 6531–6535.